Accepted Manuscript

On the weak convergence of the empirical conditional copula under a simplifying assumption

François Portier, Johan Segers

PII: S0047-259X(16)30130-0
DOI: https://doi.org/10.1016/j.jmva.2018.03.002
Reference: YJMVA 4334

To appear in: *Journal of Multivariate Analysis*

Received date : 3 November 2016

Please cite this article as: F. Portier, J. Segers, On the weak convergence of the empirical conditional copula under a simplifying assumption, *Journal of Multivariate Analysis* (2018), https://doi.org/10.1016/j.jmva.2018.03.002

On the weak convergence of the empirical conditional copula under a simplifying assumption

François Portier[a,*], Johan Segers[b]

[a] LTCI, Télécom ParisTech, Université Paris-Saclay, 46, rue Barrault, 75013 Paris, France
[b] Université catholique de Louvain, Institut de statistique, biostatistique et sciences actuarielles, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium

Abstract

A common assumption in pair-copula constructions is that the copula of the conditional distribution of two random variables given a covariate does not depend on the value of that covariate. Two conflicting intuitions arise about the best possible rate of convergence attainable by nonparametric estimators of that copula. On the one hand, the best possible rates for estimating the marginal conditional distribution functions is slower than the parametric one. On the other hand, the invariance of the conditional copula given the value of the covariate suggests the possibility of parametric convergence rates. The more optimistic intuition is shown to be correct, confirming a conjecture supported by extensive Monte Carlo simulations by Hobæk Haff and Segers [*Computational Statistics and Data Analysis* 84:1–13, 2015] and improving upon the nonparametric rate obtained theoretically by Gijbels et al. [*Scandinavian Journal of Statistics* 42:1109–1126, 2015]. The novelty of the proposed approach lies in a double smoothing procedure for the estimator of the marginal conditional distribution functions. The copula estimator itself is asymptotically equivalent to an oracle empirical copula, as if the marginal conditional distribution functions were known.

Keywords: Donsker class, empirical copula process, local linear estimator, pair-copula construction, partial copula, smoothing, weak convergence.

1. Introduction

Let (Y_1, Y_2) be a pair of continuous random variables with joint distribution function $H(y_1, y_2) = \Pr(Y_1 \le y_1, Y_2 \le y_2)$ and marginal distribution functions $F_j(y_j) = \Pr(Y_j \le y_j)$, for $y_j \in \mathbb{R}$ and $j \in \{1, 2\}$. If H is continuous, then $(F_1(Y_1), F_2(Y_2))$ is a pair of $\mathcal{U}(0, 1)$ random variables. Their joint distribution function, $D(u_1, u_2) = \Pr\{F_1(Y_1) \le u_1, F_2(Y_2) \le u_2\}$ for all $u_1, u_2 \in [0, 1]$, is therefore a copula. By Sklar's celebrated theorem [32], we have $H(y_1, y_2) = D\{F_1(y_1), F_2(y_2)\}$. The copula D thus captures the dependence between the random variables Y_1 and Y_2.

Suppose there is a third random variable, X, and suppose the joint conditional distribution of (Y_1, Y_2) given $X = x$, with $x \in \mathbb{R}$, is continuous. Then we can apply Sklar's theorem to the joint conditional distribution function $H(y_1, y_2|x) = \Pr(Y_1 \le y_1, Y_2 \le y_2|X = x)$. Writing $F_j(y_j|x) = \Pr(Y_j \le y_j|X = x)$, we have $H(y_1, y_2|x) = C\{F_1(y_1|x), F_2(y_2|x)|x\}$, where $C(u_1, u_2|x) = \Pr\{F_1(Y_1|x) \le u_1, F_2(Y_2|x) \le u_2|X = x\}$ is the copula of the conditional distribution of (Y_1, Y_2) given $X = x$. This conditional copula captures the dependence between Y_1 and Y_2 conditionally on $X = x$. Examples include exchange rates before and after the introduction of the euro [26], diastolic versus systolic blood pressure when controlling for cholesterol [22], and life expectancies of males versus females given the under-five mortality rate in a country [38].

Evidently, we can integrate out the joint and marginal conditional distributions to obtain their unconditional versions: if X has density f_X, then $H(y_1, y_2) = \int H(y_1, y_2|x)\, f_X(x)\, \mathrm{d}x$ and similarly for $F_j(y_j)$. For the copula, however, this relation does not hold: in general, $D(u_1, u_2)$ will be different from $\int C(u_1, u_2|x)\, f_X(x)\, \mathrm{d}x$.

*Corresponding author
Email addresses: francois.portier@gmail.com (François Portier), johan.segers@uclouvain.be (Johan Segers)